

378/550

FIGURE 378

MWLKVFTEFLSFATGACSGLKVTVPSTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERFHTMPKYLLGSVNKS
VVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQGNGTLSASQKIQVTVDDPVTKEPVVQIHPSPGAVEY
VGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFSPQNNTLHIAPVTKEDIGNYSCLVRNPFVSEMESDIIMPII
YYGPYGLQVNSDKGLKVGEVFTVOLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMD
YVCCAYNNITGRQDETHFTVITTSVGLEKLAQKGKSLSPASITGISLFLIISMCLLFLWKKYQPYKVIKQKLEG
RPETEYRKAQTFSGHEDALDDFGIYEFVAFFPDVSGVSRIPSRSPASDCVSGQDLHSTVYEVIQHI PAQQQDHPE

Important features:**Signal sequence:**

amino acids 1-18

Transmembrane domain:

amino acids 341-359

N-glycosylation site.

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208, 276-280, 308-312

Casein kinase II phosphorylation site.

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

Tyrosine kinase phosphorylation site.

amino acids 272-280

N-myristoylation site.

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237, 239-245

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 7-18

379/550

FIGURE 379

ATAGTAGAAGAATGTCTCTGAAATTACTGGATGAGTTTCAGTCATACTTTCACATGGGCACAA
TTCACATTCAAGCTCCTTATCCTAGGCTAATTTTATATTATGTTAAATCACTTGTTTTGT
CTCACGGCTTCCTGCCTGCTATAGGCATAATTACGAGGAAGCAGAACTTCTCCAGAAGCAAGC
GCACATGCGTTCCAAAATAAGAGCAAATTCGCTCTAAACACAGGAAAAGACCTGAAGCTTTAA
TTAAGGGGTTACATCCAACCCCAGAGCGCTTTTGTGGGCACTGATTGCTCCAGCTTCTGCGTC
ACTGCGCGAGGGAAGAGGGAAGAGGATCCAGGCGTTAGACATGTATAGACACAAAAACAGCTG
GAGATTGGGCTTAAAATACCCACCAAGCTCCAAAGAAGAGACCCAAGTCCCCAAAACATTGAT
TTCAGGGCTGCCAGGAAGGAAGAGCAGCAGCAGGGTGGGAGAGAAGCTCCAGTCAGCCCACAA
GATGCCATTGTCCCCCGGCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCACCGCTGCCCT
GCCCCTGGAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGGCAGGAATAAG
GAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGGACCTCCCAGGCCAGTGCCGG
GCCCCTCATAGGAGAGGAAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGCACCCCCCAGCA
ATCCGCGCGCCGGGACAGAATGCCCTGCAGGAACCTTCTTCTGGAAGACCTTCTCCTCCTGCAA
ATAG

380/550

FIGURE 380

MYRHKNSWRLGLKYPPSSKEETQVPKTLISGLPGRKSSSRVGEKLQSAHKMPLSPGLLLLLLS
GATATAALPLEGGPTGRDSEHMQEAAGIRKSSLLTFLAWWFEWTSQASAGPLIGEEAREVARR
QEGAPPOQSARRDRMPCRNFFWKTFSSCK

Important features:**Transmembrane domain:**

amino acids 51-69

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 35-39, 92-96

N-myristoylation sites.

amino acids 64-70, 75-81, 90-96

Amidation site.

amino acids 33-37

381/550

FIGURE 381

GGCGCCGGTGACCGGGCGGGCTGAGCGCCTCCTGCGGCCCGGCCCTGCGCGCCCCGGCCCCGCC
GCGCCGCCCACGCCCCAACCCCGGCCCGCGCCCCCTAGCCCCCGCCCGGGCCCCGCGCCCCGCGC
CCGCGCCCAGGTGAGCGCTCCGCCCCGCGCGAGGCCCGCCCCGCCCCGCCCCGCCCCGCCCC
CGGCCGCGGGGGGAACCGGGCGGATTCCTCGCGCGTCAAACCACCTGATCCCATAAAACATTG
ATCCTCCCGGCGGCCCGCGCTGCGAGCGCCCCGCCAGTCCGCGCCGCGCCGCCCTCGCCCTG
TGCGCCCTGCGCGCCCTGCGCACCCGCGGCCCGAGCCAGCCAGAGCCGGGCGGAGCGGAGCG
CGCCGAGCCTCGTCCCGCGGGCCGGGCGGGGCCGGGCGGTAGCGGCGGCGCCTGGATGCGGAC
CCGGCCGCGGGGAGACGGGCGCCCCGCCCGAAACGACTTTCAGTCCCCGACGCGCCCCGCCCA
ACCCCTACG**ATGA**AGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGGTGCTGTGGCTG
CAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGAGCCCAAGGTG
ACGACAAGCTGCCCCAGCAGGGCCTGCAGGCTGTGCCCGTGGGCATCCCTGCTGCCAGCCAG
CGCATCTTCCTGCACGGCAACCGCATCTCGCATGTGCCAGCTGCCAGCTTCCGTGCCTGCCGC
AACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCCCGAATTGATGCGGCTGCCTTCACT
GGCCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCCGGTCTGTGGACCCT
GCCACATTCACAGGCCTGGGCGCCCTACACACGCTGCACCTGGACCCTGCGGCCTGCAGGAG
CTGGGCCCCGGGGCTGTTCCGCGGCCTGGCTGCCCTGCAGTACCTCTACCTGCAGGACAACGCG
CTGCAGGCACTGCCTGATGACACCTTCCGCGACCTGGGCAACCTCACACACCTCTTCCTGCAC
GGCAACCGCATCTCCAGCGTGCCCGAGCGCGCCTTCCGTGGGCTGCACAGCCTCGACCGTCTC
CTACTGCACCAGAACCGCGTGGCCCATGTGCACCCGCATGCCTTCCGTGACCTTGGCCGCCTC
ATGACACTCTATCTGTTTGCCAACAATCTATCAGCGCTGCCCACTGAGGCCCTGGCCCCCTG
CGTGCCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGTGTGTGACTGCCGGGCACGCCCA
CTCTGGGCCTGGCTGCAGAAAGTTCCGCGGCTCCTCCTCCGAGGTGCCCTGCAGCCTCCCGCAA
CGCCTGGCTGGCCGTGACCTCAAACGCCTAGCTGCCAATGACCTGCAGGGCTGCGCTGTGGCC
ACCGGCCCTTACCATCCCATCTGGACCGGCAGGGCCACCGATGAGGAGCCGCTGGGGCTTCCC
AAGTGCTGCCAGCCAGATGCCGCTGACAAGGCCTCAGTACTGGAGCCTGGAAGACCAGCTTCG
GCAGGCAATGCGCTGAAGGGACGCGTGCCGCCCCGGTGACAGCCCGCCGGGCAACGGCTCTGGC
CCACGGCACATCAATGACTCACCCCTTGGGACTCTGCCTGGCTCTGCTGAGCCCCGCTCACT
GCAGTGCGGCCCGAGGGCTCCGAGCCACCAGGGTTCCCCACCTCGGGCCCTCGCCGGAGGCCA
GGCTGTTACGCAAGAACCGCACCCCGCAGCCACTGCCGTCTGGGCCAGGCAGGCAGCGGGGT
GGCGGGACTGGTGACTCAGAAGGCTCAGGTGCCCTACCCAGCCTCACCTGCAGCCTCACCCCC
CTGGGCCCTGGCGCTGGTGCTGTGGACAGTGCTTGGGCCCTG**TGA**CCCCCAGCGGACACAAGA
GCGTGCTCAGCAGCCAGGTGTGTGTACATACGGGTCTCTCTCCACGCCGCCAAGCCAGCCGG
GCGGCCGACCCGTGGGGCAGGCCAGGCCAGGTCTCCTGATGGACGCCTGCCGCCCGCCACC
CCATCTCCACCCCATCATGTTTACAGGGTTCGGCGGCAGCGTTTGTTCAGAACGCCGCCTC
CCACCCAGATCGCGGTATATAGAGATATGCATTTTATTTTACTTGTGTAAAAATATCGGACGA
CGTGGAATAAAGAGCTCTTTTCTTAAAAAA

382/550

FIGURE 382

MKRASAGGSRLLEAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPVGIPAASQRIF
LHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAAFGLALLEQLDLSDNAQLRSVDPATF
HGLGRLHTLHLDRCLQELGPGFLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLHGNR
ISSVPERAFRGLHSLDRLLLHQNRVAHVHPHAFRDLGRLMTLYLFANNLSALPTEALAPLRAL
QYLRRLNDNPWVCDRCRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCATATGP
YHPIWTGRATDEEPLGLPKCCQPDAAKASVLEPGRPASAGNALKGRVPPGDSPFGNGSGPRH
INDSPFGTLPGSAEPPLTAVRPEGSEPPGFPTSGPFRRRPGCSRKNRTRSHCRLGQAGSGGGGT
GDSESGALPSLTCSLTPLGLALVLWTVLGPC

Important features:**Signal peptide:**

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

383/550

FIGURE 383

TTCTGTGACCCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCCTTCTAAGAAGGGGGAGTCCTGAACTTGTCTG
AAGCCCTTGTCCGTAAGCCTTGAAC'TACGTTCTTAAATCTATGAAGTCGAGGGACCTTTCGCTGCTTTTGTAGGG
ACTTCTTTTCCTTGCTTCAGCAACATGAGGCTTTTCTTGTGGAACGCGGTCCTTGACTCTGTTCTGTCACCTTCTTTGA
TTGGGGCTTTGATCCCTGAACCAGAAAGTGAAAATTGAAGTTCTCCAGAAGCCATTCATCTGCCATCGCAAGACCA
AAGGAGGGGATTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCCTTATTTCACTCCACTCACA
AACATAACAATGGTCAGCCCATTGTTTACCCTGGGCATCCTGGAGGCTCTCAAAGGTTGGGACCAGGGCTTGA
AAGGAATGTGTGTAGGAGAGAAGAGAAAGCTCATCATTCCTCCTGCTCTGGGCTATGGAAAAGAAGGAAAAGGTA
AAATTCCCCCAGAAAGTACACTGATATTTAATATTGATCTCCTGGAGATTGGAATGGACCAAGATCCCATGAAT
CATTCCAAGAAATGGATCTTAATGATGACTGGAAACTCTCTAAAGATGAGGT'TAAAGCATATTTAAAGAAGGAGT
TTGAAAACATGGTGGGTGGTGAATGAAAGTCATCATGATGCTTTGGTGGAGGATATTTTTGATAAAGAAGATG
AAGACAAAGATGGGTTTATATCTGCCAGAGAATTTACATATAAACACGATGAGTTATAGAGATACATCTACCCCTT
TTAATATAGCACTCATCTTTCAAGAGAGGGGCAGTCATCTTTAAAGAACATTTTATTTTTATACAATGTTCTTTCT
TGCTTTGTTTTTTATTTTTATATATTTTTTCTGACTCCTATTTAAAGAACCCCTTAGGTTTCTAAGTACCCATTT
CTTTCTGATAAGTTATTGGGAAGAAAAAGCTAATTGGTCTTTGAATAGAAGACTTCTGGACAATTTTTCACTTTC
ACAGATATGAAGCTTTGTTTTACTTTCTCACTTATAAATTTAAATGTTGCAACTGGGAATATACCACGACATGA
GACCAGGTTATAGCACAAATTAGCACCCATATTTCTGCTTCCCTCTATTTTCTCCAAGTTAGAGGTCAACATTT
GAAAAGCCTTTTGCAATAGCCCCAAGGCTTGCTATTTTCATGTTATAATGAAATAGTTTATGTGTAAGTGGCTCTG
AGTCTCTGCTTGAGGACCAGAGGAAAATGGTTGT'TGGACCTGACTTGTTAATGGCTACTGCTTTACTAAGGAGAT
GTGCAATGCTGAAGTTAGAAAACAAGGTTAATAGCCAGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGA
GGCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTCGAGACCAGCCTGACCAACACGGAGAAACCCTATCTCTAC
TAAAAATACAAAGTAGCCCGCGGTGCTGATGCGTGCCTGTAATCCCAGCTACCCAGGAAGGCTGAGGCGGCAGAA
TCACTTGAACCCGAGGCGGAGGTTGCGGTAAGCCGAGATCACCTNCAGCCTGGACACTCTGTCTCGAAAAAAGAA
AAGAACACGGTTAATACCATATNAATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAGCTTGG
CTCCTAGTGATTGGTGGCCTATTATGATAAATAGGACAAATCATTTATGTGTGAGTTTCTTTGTAATAAAATGTA
TCAATATGTTATAGATGAGGTAGAAAGTTATATTTATATTCAATATTTACTTCTTAAGGCTAGCGGAATATCCTT
CCTGGTTCTTTAATGGGTAGTCTATAGTATATTATACTACAATAACATTGTATCATAAGATAAAGTAGTAACCA
GTCTACATTTTCCCATTTCTGTCTCATCAAAAAC'TGAAGTTAGCTGGGTGTGGTGGCTCATGCCTGTAATCCCAG
CACTTTGGGGGCCAAGGAGGTTGGATCACTTGAGATCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCT
TGTCTCTACTAAAAATACAAAAATTAGCCAGGCGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAGGCTGA
GACAGGAGATTTGCTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATTGTGCCACTGCACTCCAGCCTGGG
TGACAGAGCAAGACTCCATCTCAAAAAAAAAAAAAAGGAAGCAGACCTACAGCAGCTACTATTGAATAAATACCTA
TCCTGGATTTT

384/550

FIGURE 384

MRLFLWNAVLTLFVTSLLIGALIPPEVKIEVLQKPFICHRKTKGGDLMLVHYEGYLEKDGSLF
HSTHKHNNQPIWFTLGILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPESTLI
FNIDLLAIRNGPRSHESFQEMDLNDDWKLSKDEVKAYLKKEFEKHGAVVNESHHDALVEDIFD
KEDEDKDGFISAREFTYKHDEL

Important features:**Signal peptide:**

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

385/550

FIGURE 385

CTCCCACGGTGTCCAGCGCCCAGAAATGCGGCTTCTGGTCTGCTATGGGGTTGCCTGCTGCTC
CCAGGTTATGAAGCCCTGGAGGGGCCAGAGGAAATCAGCGGGTTCGAAGGGGACACTGTGTCC
CTGCAGTGCACCTACAGGGAAGAGCTGAGGGACCACCGGAAGTACTGGTGCAGGAAGGGTGGG
ATCCTCTTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAATGAAG
GGCAGGGTGTCCATCCGTGACAGCCGCCAGGAGCTCTCGCTCATTGTGACCCTGTGGAACCTC
ACCCTGCAAGACGCTGGGGAGTACTGGTGTGGGGTCGAAAAACGGGGCCCCGATGAGTCTTTA
CTGATCTCTCTGTTCGTCTTTCCAGGACCCTGCTGTCTCTCCCTCCCCTTCTCCACCTTCCAG
CCTCTGGCTACAACACGCCTGCAGCCCAAGGCAAAAGCTCAGCAAACCCAGCCCCCAGGATTG
ACTTCTCCTGGGCTCTACCCGGCAGCCACCACAGCCAAGCAGGGGAAGACAGGGGCTGAGGCC
CCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACACAGGAACCTCT
CCTCACCCAGCGACCTCTCCTCCTGCAGGGAGCTCCCGCCCCCCCCATGCAGCTGGACTCCACC
TCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGGGTGTCCATCCCG
ATGGTCCGCATACTGGCCCCAGTCTGGTGTCTGCTGAGCCTTCTGTGAGCCGCAGGCCTGATC
GCCTTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTCAACAGGCCACGGAGACACAGAGG
AACGAGAAGTTCTGGCTCTCACGCTTGACTGCGGAGGAAAAGGAAGCCCCCTTCCCAGGCCCT
GAGGGGACGTGATCTCGATGCCTCCCCTCCACACATCTGAGGAGGAGCTGGGCTTCTCGAAG
TTTGTCTCAGCGTAGGGCAGGAGGCCCTCCTGGCCAGGCCAGCAGTGAAGCAGTATGGCTGGC
TGGATCAGCACCGATTCCCGAAAGCTTTCACCTCAGCCTCAGAGTCCAGCTGCCCGGACTCC
AGGGCTCTCCCCACCTCCCCAGGCTCTCCTCTTGATGTTCAGCCTGACCTAGAAGCGTTT
GTCAGCCCTGGAGCCCAGAGCGGTGGCCTTGCTCTTCCGGCTGGAGACTGGGACATCCCTGAT
AGGTTACATCCCTGGGCAGAGTACCAGGCTGCTGACCCTCAGCAGGGCCAGACAAGGCTCAG
TGGATCTGGTCTGAGTTTCAATCTGCCAGGAACTCCTGGGCCTCATGCCAGTGTGCGACCCT
GCCTTCTCTCCACTCCAGACCCACCTTGTCTTCCCTCCCCTGGCGTCTCAGACTTAGTCCCA
CGGTCTCCTGCATCAGCTGGTGTGATGAAGAGGAGCATGCTGGGGTGAGACTGGGATTCTGGCTT
CTCTTTGAACCACCTGCATCCAGCCCTTCAGGAAGCCTGTGAAAACGTGATTCTGGCCCCA
CCAAGACCCACCAAAACCATCTCTGGGCTTGGTGCAGGACTCTGAATTCTAACAATGCCAGT
GACTGTGCACTTGAGTTTGAGGGCCAGTGGGCCTGATGAACGCTCACACCCCTTCAGCTTAG
AGTCTGCATTTGGGCTGTGACGTCTCCACCTGCCCAATAGATCTGCTCTGTCTGCGACACCA
GATCCACGTGGGGACTCCCCTGAGGCCTGCTAAGTCCAGGCCTTGGTCAAGTCAAGTGCACAT
TGCAGGATAAGCCCAGGACCGGCACAGAAGTGGTTGCCTTTNCCATTTGCCCTCCCTGGNCCA
TGCCCTTCTTGCCCTTGGA AAAAATGATGAAGAAAACCTTGGCTCCTTCTTGTCTGGAAAGGG
TTACTTGCCTATGGGTCTGGTGGCTAGAGAGAAAAGTAGAAAACCAGAGTGCACGTAGGTGT
CTAACACAGAGGAGAGTAGGAACAGGGCGGATACCTGAAGGTGACTCCGAGTCCAGCCCCCTG
GAGAAGGGGTCGGGGGTGGTGGTAAAGTAGCACAACCTACTATTTTTTTTCTTTTCCATTATT
ATTGTTTTTTAAGACAGAATCTCGTGCTGCTGCCAGGCTGGAGTGCAGTGGCACGATCTGCA
AACTCCGCCTCCTGGGTTCAGTGATTTCTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAG
GCACGCACCACACACCTGGCTAATTTTTTGTACTTTTAGTAGAGATGGGGTTTTACCATGTTG
GCCAGGCTGGTCTTGAACCTCCTGACCTCAAATGAGCCTCCTGCTTCAGTCTCCCAAATTGCCG
GGATTACAGGCATGAGCCACTGTGTCTGGCCCTATTTCTTTAAAAAGTGAAATTAAGAGTTG
TTCAGTATGCAAACTTGGAAGATGGAGGAGAAAAAGAAAAGGAAGAAAAAATGTCACCCA
TAGTCTCACCAGAGACTATCATTATTTCTGTTTGTGTTGTTACTTCTTCCACTCTTTCTTCTTC
ACATAATTTGCCGGTGTCTTTTTACAGAGCAATTATCTTGTATATACAACCTTTGTATCCTGC
CTTTTCCACCTTATCGTTCCATCACTTTATTCCAGCACTTCTCTGTGTTTTACAGACCTTTTT
ATAAATAAAATGTTTCATCAGCTGCATAAAAAAAAAAAAAA

386/550

FIGURE 386

MRLLVLLWGCLLLPGYEALEGP EEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCSG
TIYAE EEGQETMKGRVSI RDSRQELSLIVTLWNLT LQDAGEYWC GVEKRGPD E SLLISLFVFP
GPCCPPSPSP T FQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPGTSQ
YGHERTSQYTGTSPHPATSPPAGSSRPPMQLDSTSAEDTSPALSSGSSKPRVSIPMVRILAPV
LVLLSLLSAAGLIAFCSHLLLWRKEAQQATETQRNEKFWLSRLTAE E KEAPSQAPEGDVISM P
PLHTSEEELGFSKFVSA

Important features:**Signal peptide:**

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

387/550

FIGURE 387

GCGCCGGGAGCCCATCTGCCCCAGGGGCACGGGGCGCGGGGCCGGCTCCCGCCCGGCACATG
GCTGCAGCCACCTCGCGCGCACCCCGAGGCGCCGCGCCAGCTCGCCCGAGGTCCGTCCGAGG
CGCCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGCGCCCGCGTCCGGGGATCGGG
ATGTCCCTCCTCCTTCTCCTCTTGCTAGTTTCCTACTATGTTGGAACCTTGGGGACTCACACT
GAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTTGGCCTGCCACCATCAACTGGGGCTTCCA
GAAAAAGACACTCTGGATATTGAATGGCTGCTCACCGATAATGAAGGGAACCAAAAAGTGGTG
ATCACTTACTCCAGTCGTCTATGTCTACAATAACTTGACTGAGGAACAGAAGGGCCGAGTGGCC
TTTGCTTCCAATTTCTTGGCAGGAGATGCCTCCTTGAGATTGAACCTCTGAAGCCCAGTGAT
GAGGGCCGGTACACCTGTAAGGTTAAGAATTCAGGGCGCTACGTGTGGAGCCATGTCATCTTA
AAAGTCTTAGTGAGACCATCCAAGCCCAAGTGTGAGTTGGAAGGAGAGCTGACAGAAGGAAGT
GACCTGACTTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATTACTGGCAGCGA
ATCCGAGAGAAAGAGGGAGAGGATGAACGTCTGCCTCCCAAATCTAGGATTGACTACAACCAC
CCTGGACGAGTTCTGCTGCAGAATCTTACCATGTCTACTCTGGACTGTACCAGTGCACAGCA
GGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGCAGTAAGTGTACAGTATGTACAAAGCATC
GGCATGGTTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTGATTTTCCTCTTGGTG
TGGCTGCTAATCCGAAGGAAAGACAAAGAAAGATATGAGGAAGAAGAGAGACCTAATGAAATT
CGAGAAGATGCTGAAGCTCCAAAAGCCCGTCTTGTGAAACCCAGCTCCTCTTCCTCAGGCTCT
CGGAGCTCACGCTCTGGTTCTTCTCCACTCGCTCCACAGCAAATAGTGCCTCACGCAGCCAG
CGGACACTGTCAACTGACGCAGCACCCAGCCAGGGCTGGCCACCCAGGCATACAGCCTAGTG
GGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATGCTAATCTGACCAAAGCAGAA
ACCACACCCAGCATGATCCCCAGCCAGAGCAGAGCCTTCCAAACGGTCT**TGA**ATTACAATGGAC
TTGACTCCCACGCTTTCCTAGGAGTCAGGGTCTTTGGACTCTTCTCGTCATTGGAGCTCAAGT
CACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCAGTGAGCATTGCACGGAACAGATT
CAGATGAGCATTTTTCCTTATACAATACCAAACAAGCAAAAGGATGTAAGCTGATTCATCTGTA
AAAAGGCATCTTATTGTGCCTTTAGACCAGAGTAAGGGAAAGCAGGAGTCCAAATCTATTTGT
TGACCAGGACCTGTGGTGAGAAGGTTGGGGAAAGGTGAGGTGAATATACCTAAACTTTTAAT
GTGGGATATTTTGTATCAGTGCTTTGATTACAAATTTTCAAGAGGAAATGGGATGCTGTTTGT
AAATTTTCTATGCATTTCTGCAAACCTTATTGGATTATTAGTTATTCAGACAGTCAAGCAGAAC
CCACAGCCTTATTACACCTGTCTACACCATGTACTGAGCTAACCCTTCTAAGAACTCCAAA
AAAGGAAACATGTGTCTTCTATTCTGACTTAACCTTCATTTGTCATAAGGTTTGGATATTAATT
TCAAGGGGAGTTGAAATAGTGGGAGATGGAGAAGAGTGAATGAGTTTCTCCCACTCTATACTA
ATCTCACTATTTGTATTGAGCCCAAAATAACTATGAAAGGAGACAAAATTTGTGACAAAGGA
TTGTGAAGAGCTTTCATCTTCATGATGTTATGAGGATTGTTGACAAACATTAGAAATATATA
ATGGAGCAATTGTGGATTTCCCCTCAAATCAGATGCCTCTAAGGACTTTCCTGCTAGATATTT
CTGGAAGGAGAAAATACAACATGTCAATTTATCAACGTCCTTAGAAAGAATTCTTCTAGAGAAA
AAGGGATCTAGGAATGCTGAAAGATTACCCAACATAACCATTATAGTCTCTTCTTCTGAGAAA
ATGTGAAACCAGAATTGCAAGACTGGGTGGACTAGAAAGGGAGATTAGATCAGTTTCTCTTA
ATATGTCAAGGAAGGTAGCCGGGCATGGTGCCAGGCACCTGTAGGAAAATCCAGCAGGTGGAG
GTTGCAGTGAGCCGAGATTATGCCATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

388/550

FIGURE 388

MSLLLLLLLLVSYYVGTLLGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLLTDNEGNQKV
ITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHVIL
KVLVRPSKPKCELEGELTEGSDTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRIDYNH
PGRVLLQNLTMSYSGLYQCTAGNEAGKESCVVRVTVQYVQSIGMVAGAVTGIVAGALLIFLLV
WLLIRRKDKERYEEEEERPNEIREDAEAPKARLVKPSSSSSGSRSSRSGSSSTRSTANSASRSQ
RTLSTDAAPQPGLATQAYSLVGPEVRGSEPKKVHHANLTKAETTPSMIPSQSRAFAQTV

Important features:**Signal sequence:**

amino acids 1-16

Transmembrane domain:

amino acids 232-251

389/550

FIGURE 389

GCGGCACCTGGAAGATGCGCCCATTTGGCTGGTGGCCTGCTCAAGGTGGTGTTCGTGGTCTTCG
CCTCCTTGTGTGCCTGGTATTCGGGGTACCTGCTCGCAGAGCTCATTCCAGATGCACCCCTGT
CCAGTGCTGCCTATAGCATCCGCAGCATCGGGGAGAGGCCTGTCTCAAAGCTCCAGTCCCCA
AAAGGCCAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCTACAGGTTACTCA
GCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACCTACTTATGGGAGAAC
AGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACTATGTAAGTGGGAATGTGA
CAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTGGACCGATGACAAAGTTTATTC
AGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCTATGACGACGGAAGCACAAGACTGA
ATAACGATGCCAAGAATGCCATAGAAGCACTTGGAAGTAAAGAAATCAGGAACATGAAATTCA
GGTCTAGCTGGGTATTTATTGCAGCAAAAGGCTTGGAAGTCCCTTCCGAAATTCAGAGAGAAA
AGATCAACCACTCTGATGCTAAGAACAACAGATATTCTGGCTGGCCTGCAGAGATCCAGATAG
AAGGCTGCATACCCAAAGAACGAAGCTTGACACTGCAGGGTCCTGAGTAAATGTGTTCTGTATA
AACAAATGCAGCTGGAATCGCTCAAGAATCTTATTTTTCTAAATCCAACAGCCCATATTTGAT
GAGTATTTTGGGTTTCTTGTAACCAATGAACATTTGCTAGTTGTATCAAATCTTGGTACGCA
GTATTTTATACCAGTATTTTATGTAGTGAAGATGTCAATTAGCAGGAACTAAAATGAATGG
AAATTCTTAAAAAAAAA

390/550

FIGURE 390

MRPLAGGLLKVVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPKRQKC
DHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTATRC
FDMYEGDNSGPMTKFIQSAAFKSLLEFMTYDDGSTRLNNDKNAIEALGSKEIRNMKFRSSWV
FIAAKGLELPSEIQREKINHSDAKNNRYSGWPAEIQIEGCIPKERS

Important features:**Signal sequence.**

amino acids 1-20

N-glycosylation sites.

amino acids 120-124, 208-212

Glycosaminoglycan attachment site.

amino acids 80-84

N-myristoylation sites.

amino acids 81-87, 108-114, 119-125

392/550

FIGURE 392

MSRPGTATPALALVLLAVTLAGVGAQGALEDPDYYGQEIWSREPYYARPEPELETFSPPLPA
GPGEWEERRPQEPRPPKRATKPKKAPKREKSAPEPPPPGKHSNKKVMRTKSSEKAANDDHVR
VAREDVRESCFPLGLETLKITDFQLHASTVKRYGLGAHRGRLNIQAGINENDFYDGAWCAGR
DLQQWIEVDARRLTRFTGVITQGRNSLWLSDWVTSYKVMVSNDSTWVTVKNGSGDMIFEGNS
EKEIPVLNELPVPMVARYIRINPQSWFDNGSICMRMEILGCPLPDPNNYYHRRNEMTTTDDL
FKHHNYKEMRQLMKVVNEMCPNITRIYNIGKSHQGLKLYAVEISDHPGEHEVGEPEFHYIAGA
HGNEVLGRELLLLLVQFVCQEYLARNARIVHLVEETRIHVLPSLNPDGYEKAYEGGSELGGWS
LGRWTHDGDINNNFPDLNTLLWEAEDRQNVPRKVPNHYIAIPEWFLSENATVAAETRAVIAW
MEKIPFVLGGNLQGGELVVAYPYDLVRSPWKTQEHTPTPDDHVFRWLAYSASTHRLMTDARR
RVCHTEDFQKEEGTVNGASWHTVAGSLNDFSYLHTNCFELSIYVGCDKYPHESQLPEEWENNR
ESLIVFMEQVHRGIKGLVRDSHGKGI PNAIISVEGINHDIRTANDGDYWRLNPGEYVVTAKA
EGFTASTKNCMVGYDMGATRCDFTLSTKNMARIREIMEKFGKQPVSLPARRLKLRGRKRRQRG

393/550

FIGURE 393

GTCCACATCCTGCTCAACTGGGTCAAGTCCCTCTTAGACCAGCTCTTGTCCATCATTGCTGAAGTGGACCAAC
 TAGTTCCCCAGTAGGGGGTCTCCCTGGCAATTCTTGATCGGGCTTTGGACATCTCAGATCGCTTCCAATGAAGA
 TGGCCTTGCTTGGGGTCTGCTTGTTCATAATCATCTAACTATGGGACAAGGTTGTGCCGGCAGCTCTGGGGG
 AAGGAGCACGGGGCTGATCAAGCCATCCAGGAACACTGGAGGACTTGTCCAGCCTTGAAAGAAGCTCTAGTGGTT
 TCTGAATCTAGCCCACTTGGCGGTAAGCAATGATGCAACTTCTGCAACTTCTGCTGGGGCTTTTGGGGCCAGGTGG
 CTACTTATTTCTTTAGGGGATTGTCAAGGAGGTGACCACCTCTCACGGTGAAATACCAAGTGTCAAGGAAAGTGCC
 ATCTGGTACAGTGATCGGGAAGCTGTCCCAGGAAGTGGGCCGGGAGGAGAGGCGGAGGCAAGCTGGGGCCGCTT
 CCAGGTGTTGCAGCTGCCTCAGGCGCTCCCCATTAGGTGGACTCTGAGGAAGGCTTGCTCAGCACAGGCAGGCG
 GCTGGATCGAGAGCAGCTGTGCCGACAGTGGGATCCCTGCCTGGTTTCTTTGATGTGCTTGCCACAGGGGATTT
 GGCTGATCCATGTGGAGATCCAAGTGTGACATCAATGACCACAGCCAGGTTTCCCAAAGGCGAGCAGGA
 GTGGAAATCTCTGAGAGCGCTCTCTGCGAACCCTGGATCCCCCTGGACAGAGCTCTTGACCCAGACACAGGCC
 TAACACCCTGCACACCTACACTCTGTCTCCAGTGAGCACTTTGCCTTGGATGTCTATTGTGGGCCCTGATGAGAC
 CAAACATGCAGAAGCTCATAGTGGTGAAGGAGCTGGACAGGGAAATCCATTCAATTTTTTGATCTGGTGTAACTGC
 CTATGACAATGGGAACCCCCCAAGTCAAGTACCAGCTTGGTCAAGGTCAACGCTTGGACTCCAATGACAATAG
 CCCTGCGTTTGTGAGAGTTCACTGGCACTGGAAATCCAAGAAGATGCTGCACCTGGTACGCTTCTCATAAACT
 GACCGCCACAGACCCTGACCAAGGCCCAATGGGGAGGTGGAGTTCTTCTCAGTAAGCACATGCCTCCAGAGGT
 GCTGGACACCTTCAGTATTGATGCCAAGACAGGCCAGGTCAATCTGCGTCGACCTCTAGACTATGAAAAGAACC
 TGCCCTACGAGGTGGATGTTTCAAGCAAGGGACCTGGGTCCCAATCCTATCCAGCCCATTCGAAAGTTCTCATCAA
 GGTCTGGATGTCAATGACAACATCCCAAGCATCCACGTCAACATGGGCTTCCAGCCATCACTGGTGTGAGAAGC
 TCTTCCCAAGGACAGTTTTTATTGCTCTTGTCTATGGCAGATGACTTGGATTCAAGGACACAATGGTTTGGTCCACTG
 CTGGCTGAGCCAAGAGCTGGGCCACTTCAGGCTGAAAAGAACTAATGGCAACACATACATGTTGCTAACCAATGC
 CACACTGGACAGAGAGCAGTGGCCCAATATACCCTCACTCTGTAGCCCAAGACCAAGGACTCCAGCCCTTATC
 AGCCAAGAAACAGCTCAGCATTGATGATGACATCAACGACAATGCACCTGTGTTGAGAAAAGCAGGTATGA
 AGTCTCCACGCGGGAAACAACCTTACCCTCTCTTACCTCATTACCATCAAGGCTCATGATGCAGACTTGGGCAT
 TAATGGAAAAGTCTCATACCGCATCCAGGACTCCCCAGTTGCTCACCTTAGTAGCTATTGACTCCAACACAGGAGA
 GGTCACTGCTCAGAGGTCACTGAACATGAAGAGATGGCCGGCTTTGAGTCCAGGTGATCGCAGAGGACAGCGG
 GCAACCCATGCTTGCATCCAGTGTCTCTGTGTGGGTCAAGCTCTTGGATGCCAATGATAATGCCCCAGAGGTGTT
 CCAGCCTGTGCTCAGCGATGGAAAAGCCAGCCTCTCCGTGCTTGTGAATGCCCTCCACAGGCCACCTGCTGGTGCC
 CATCGAGACTCCCAATGGCTTGGGCCAGCGGGCACTGACACACCTCCACTGGCCACTCACAGCTCCCGGCCATTT
 CCTTTTGACAACCAATTGTGGCAAGAGATGCAGACTCGGGGGCAATGGAGAGCCCCCTTACAGCATCCGCAATGG
 AAATGAAGCCACCTCTTCATCCTCAACCCTCATACGGGGCAGCTGTTGCTCAATGTACCAATGCCAGCAGCCT
 CATTGGGAGTGAGTGGGAGCTGGAGATAGTAGTAGAGGACCAGGGAAGCCCCCTTACAGACCCGAGCCCTGTT
 GAGGCTCATGTTTGTACACAGTGTGGACCACCTGAGGGACTCAGCCCGCAAGCCTGGGGCTTTGAGCATGTGAT
 GCTGACGGTGATCTGCCTGGCTGTACTGTTGGGCATCTTGGGTGATCTTGGCTTTGTTTATGTCTCATCTGCCG
 GACAGAAAAGAAAGGACAACAGGGCCTACAACCTGTCCGGAGGCGAGTCCACCTACCGCCAGCAGCCCAAGAGGCC
 CCAGAAACAGATTGAGAAGGACAGATCCACCTGCTGCTGTGCTCAGGGCTCAGGCAGGTGAGCTTGTGAAGT
 CGGGCAGTCCACAAAGATGTGGACAAGGAGGCGATGATGGAAGCAGGCTGGGACCCCTGCCCTGCAGGCCCTTT
 CCACCTCAACCCGACCTGTACAGGACGCTGCGTAATCAAGGCAACCAGGAGCACCAGCGGAGAGCCGAGAGGT
 GCTGCAAGACACGGTCAACCTCCTTTTCAACCATCCAGGCAGAGGAATGCCCTCCCGGAGAACCTGAACCTTCC
 CGAGCCCGAGCCTGCCACAGGCCAGCCACGTTCCAGGCCTCTGAAGGTTGCAGGCAGCCCCACAGGGAGGCTGGC
 TGGAGACCAGGGCAGTGAGGAAGCCCCACAGAGGCCACCAGCCTCCTCTGCAACCTGAGACGGCAGCGACATCT
 CAATGGCAAAGTGTCCCTGAGAAAGAATCAGGGCCCCGTGAGATCCTGCGGAGCCTGGTCCSGCTGTCTGTGGC
 TGCCCTTCCCGAGCGGAACCCCGTGGAGGAGCTCACTGTGATTCTCCTCCTGTTTACGCAATCTCCAGCTGCT
 GTCTTGTGCTCAGGGCCAATTCAGCCCAACCAACCCAGGAAATAGTACTTGGCCAAGCCAGGAGG
 CAGCAGGAGTGCAATCCAGACACAGATGGCCCAAGTGCAAGGGCTGGAGGCCAGACAGACCCAGAACAGGAGGA
 AGGGCCTTTGGATCCTGAAGAGGACCTCTCTGTGAAGCAACTGCTAGAAGAAGAGCTGTCAAGTCTGCTGGACCC
 CAGCACAGGTCTGGCCCTGGACCGGCTGAGCGCCCTGACCCGGCTGGATGGCGAGACTCTCTTTGCCCTCAC
 CACCAACTACCGTGACAATGTGATCTCCCGGATGCTGCAGCCACGGAGGAGCCGAGGACCTTCCAGACGTTCCG
 CAAGGCAGAGGCACAGAGCTGAGCCCAACAGGCACGAGGCTGGCCAGCACCTTTGCTCTCGGAGATGAGCTCACT
 GCTGGAGATGCTGCTGGAACAGCGCTCCAGCATGCCCGTGGAGGCCGCTCCGAGGCGCTGCGGCGGCTCTCGGT
 CTGCGGGAGGACCCTCAGTTTAGACTTGGCCACAGTGCAGCCTCAGGCATGAAAGTGCAAGGGGACCCAGGTGG
 AAAGACGGGGACTGAGGGCAAGAGCAGAGGCAGCAGCAGCAGCAGGTGCTGTGAACATACTCAGACCGCT
 CTGGATCCAAGAACCAGGGGCTGAGGATCTGTGGACAAGAGCTGGTTTCTAAATCTTGTAACTCACTAGCTAG
 CGGCGCCTGAGAACCTTTAGGGTGACTGATGCTACCCCCACAGAGGAGGCAAGAGCCCCAGGACTAACAGCTGAC
 TGACCAAAGCAGCCCTTGTAAAGCAGCTCTGAGTCTTTTGGAGGACAGGCACGGTTTGTGGCTGAGATAAGTGT
 TCCTGGCAAACATATGTGGAGCACAAAGGGTCACTCCTCTGGCAGAACAGATGCCACGGAGTATCACAGGCAGG
 AAAGGGTGGCCTTCTTGGGTAGCAGGAGTCAAGGGGCTGTACCTTGGGGTGGCAGGAAATGCTCTGTGACCTAT
 CAATAAAGGAAAAGCAGTAAAAA

394/550

FIGURE 394

MMQLLQLLLGLLGPGGYLFLLGDCQEVTTTLTVKYQVSEEVPSGTVIGKLSQELGREERRRQAG
 AAFQVLQLEQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSEFDVLATGDLALIHVEIQVL
 DINDHQPFRFPKGEQELEISESASLRTRIPLDRALDPDTGPNLTHTYTLSPSEHFALDVIVGPD
 ETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKSGTSLVKVNVLDSDNDNSPFAFAESSLALE
 IQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPPEVLDTFSIDAKTGQVILRRPLDYEKN
 PAYEVDVQARDLGPNPIPAHCKVLIKVLVDVNDNIPSIHVTWASQPSLVSEALPKDSFIALVMA
 DDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQDQGLQPLS
 AKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHDADLGINGKVSRYRIQDSPV
 AHLVAIDSNTGEVTAQRSLNVEEMAGFEFQVIAEDSGQPMLASSVSVWVSLLDANDNAPEVVQ
 PVLSDGKASLSVLVNASSTGHLLVPIETPNGLGPAGTDTPLATHSSRPFLTTIVARDADSGA
 NGEFLYSIRNGNEAHLFILNPHTGQLFVNVTNASSLIGSEWELEIIVVEDQGSFPLQTRALLRV
 MFVTSVDHLRDSARKPGALSMSMLTVICLAVLLGIFGLILALFMSICRTEKKDNRAYNCREAE
 STYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEAGWDPCLOAPFHILT
 PTLYRTLNRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLPPEQPATGQPRSRPLKV
 AGSPTGRLAGDQGSEEPQRPPASSATLRRQRHLNGKVSPEKESGPRQILRSLVRLSVAFAE
 RNPVEELTVDSPPVQQISQLLSLLHQGFQPKPNHRGNKYLAKEGGSRSAIPD TDGPSARAGG
 QTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPDPAWMARLSLPLTTNYRD
 NVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSLLEMLLEQRSSMPVEAASE
 ALRRLSVCGRTLSLDLATSAASGMKVQGDPPGGKTGTEGKSRGSSSSSRCL

Important features:**Signal peptide:**

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

395/550

FIGURE 395

CCCAGGCTCTAGTGCAGGAGGAGAAGGAGGAGGAGCAGGAGGTGGAGATTCCCAGTTAAAAGG
CTCCAGAATCGTGTACCAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAATC
AGTAGGTGACCCCGCCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCCCGACCTCGTGCG
GCCAAGACGTGGATGTTCTGCTCTTGCTGGGGGGAGCCTGGGCAGGACACTCCAGGGCACAG
GAGGACAAGGTGCTGGGGGGTTCATGAGTGCCAACCCCATTCGCAGCCTTGGCAGGCGGCCCTTG
TTCCAGGGCCAGCAACTACTCTGTGGCGGTGTCCTTGTAGGTGGCAACTGGGTCCTTACAGCT
GCCCCTGTAAAAAACCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAATAAAGAT
GGCCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCCTGCTACAACAGCAGCGAT
GTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTGCGTGACCAGGCATCCCTGGGGTCC
AAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCA
GGCTGGGGCACTGTCACCAGTCCCCGAGAGAATTTTCCTGACACTCTCAACTGTGCAGAAGTA
AAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCACAGATGGCATGGTC
TGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGGGCGATTCTGGAGGCCCCCTGGTGTGT
GATGGTGCCTCCAGGGCATCACATCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCT
GGCGTCTATACCAACATCTGCCGCTACCTGGACTGGATCAAGAAGATCATAGGCAGCAAGGGC
TGATTCTAGGATAAGCACTAGATCTCCCTTAATAAACTCACAACTCTCTGGTTC

396/550

FIGURE 396

MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECQPHSQPWQAALFQGQQLLCGGVLV
GGNWVLTAAHCKKPKYTVRLGDHSLQNKDGPEQEIPVVQSIHPHCYNSSDVEDHNNHDLMLLQL
RDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCEDAYP
GQITDGMVCAGSSKGADTCQGDSSGGLVCDGALQGITSWGS DPCGRSDKPGVYTNICRYLDWI
KKIIGSKG

Important Features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

397/550

FIGURE 397

GGCGGCTGCTGAGCTGCCTTGAGGTGCAGTGTTGGGGATCCAGAGCCATGTCGGACCTGCTAC
TACTGGGCCTGATTGGGGGCCTGACTCTCTTACTGCTGCTGACGCTGCTGGCCTTTGCCGGGT
ACTCAGGGCTACTGGCTGGGGTGGAAGTGAGTGCTGGGTACCCCCCATCCGCAACGTCACTG
TGGCCTACAAGTTCCACATGGGGCTCTATGGTGAGACTGGGCGGCTTTTCACTGAGAGCTGCA
GCATCTCTCCCAAGCTCCGCTCCATCGCTGTCTACTATGACAACCCCCACATGGTGCCCCCTG
ATAAGTGCCGATGTGCCGTGGGCAGCATCCTGAGTGAAGGTGAGGAATCGCCCTCCCCTGAGC
TCATCGACCTCTACCAGAAATTTGGCTTCAAGGTGTTCTCCTTCCCGGCACCCAGCCATGTGG
TGACAGCCACCTTCCCCTACACCACCATTTCTGTCCATCTGGCTGGCTACCCGCCGTGTCCATC
CTGCCTTGGACACCTACATCAAGGAGCGGAAGCTGTGTGCCTATCCTCGGCTGGAGATCTACC
AGGAAGACCAGATCCATTTTCATGTGCCACTGGCACGGCAGGGAGACTTCTATGTGCCTGAGA
TGAAGGAGACAGAGTGGAATGGCGGGGGCTTGTGGAGGCCATTGACACCCAGGTGGATGGCA
CAGGAGCTGACACAATGAGTGACACGAGTTCTGTAAGCTTGGAAGTGAGCCCTGGCAGCCGGG
AGACTTCAGCTGCCACACTGTCACCTGGGGCGAGCAGCCGTGGCTGGGATGACGGTGACACCC
GCAGCGAGCACAGCTACAGCGAGTCAGGTGCCAGCGGCTCCTCTTTTGAGGAGCTGGACTTGG
AGGGCGAGGGGCCCTTAGGGGAGTCACGGCTGGACCTGGGACTGAGCCCCCTGGGGACTACCA
AGTGGCTCTGGGAGCCCCTGCCCCTGAGAAGGGCAAGGAGTAACCCATGGCCTGCACCCCTCC
TGCAGTGCAGTTGCTGAGGAACTGAGCAGACTCTCCAGCAGACTCTCCAGCCCTCTTCCTCCT
TCCTCTGGGGGAGGAGGGGTTCTTGAGGGACCTGACTTCCCCTGCTCCAGGCCTCTTGCTAAG
CCTTCTCCTCACTGCCCTTTAGGCTCCCAGGGCCAGAGGAGCCAGGGACTATTTTCTGCACCA
GCCCCAGGGCTGCCGCCCTGTTGTGTCTTTTTTTTCAGACTCACAGTGGAGCTTCCAGGACC
CAGAATAAAGCCAATGATTTACTTGTTTCACCTGGAAAAAAAAAAAAAAAAA

398/550

FIGURE 398

MSDLLLLGLIGGLTLLLLLTLLAFAGYSGLLAGVEVSAGSPPIRNVTVAYKFHMGLYGETGRL
FTESCSISPKLRSIAVYYDNPHMVPPDKCRCVGSILSEGEESPSPELIDLYQKFGFKVFSFP
APSHVVTATFPYTTILSIWLATTRVHPALDTYIKERKLCAYPRLEIYQEDQIHMCPLARQGD
FYVPEMKETEWKWRGLVEAIDTQVDGTGADTMSDTSSVSLEVSPGSRETSAAATLSPGASSRGW
DDGDTRSEHSYSESGASGSSFEELDLEGEGLGESRLDPGTEPLGTTKWLWEPTAPEKGKE

400/550

FIGURE 400

MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFNLRTSKDPEHEG
CYLSVGHSQPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLVSALHTREKDANVVVVDWLPL
AHQLYTDVNNTRVVGHSIARMLDWLQEKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGTVGRI
TGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIYPNGGDFQPGC
GLNDVLGSIAYGTITEVVKCEHERAVHLFVDSL VNQDKPSFAFQCTDSNRFKKGICLSCRKNR
CNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP

Important features:**Signal peptide:**

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

401/550

FIGURE 401

CTTCCCAGCCCTGTGCCCCAAAGCACCTGGAGCATATAGCCTTGCAGAACTTCTACTTGCCTG
CCTCCCTGCCTCTGGCC**ATG**GCCTGCCGGTGCCTCAGCTTCCTTCTGATGGGGACCTTCCTGT
CAGTTTCCCAGACAGTCCTGGCCCAGCTGGATGCACTGCTGGTCTTCCCAGGCCAAGTGGCTC
AACTCTCCTGCACGCTCAGCCCCCAGCACGTCAACATCAGGGACTACGGTGTGTCCTGGTACC
AGCAGCGGGCAGGCAGTGCCCCCTCGATATCTCCTCTACTACCGCTCGGAGGAGGATCACCACC
GGCCTGCTGACATCCCCGATCGATTCTCGGCAGCCAAGGATGAGGCCACAAATGCCTGTGTCC
TCACCATTAGTCCCGTGCAGCCTGAAGACGACGCGGATTACTACTGCTCTGTTGGCTACGGCT
Ttagtccc**TAG**GGGTGGGGTGTGAGATGGGTGCCTCCCCTCTGCCTCCCATTTCTGCCCCCTGA
CCTTGGGTCCCTTTTAACTTTCTCTGAGCCTTGCTTCCCCTCTGTAAAATGGGTTAATAATA
TTCAACATGTCAACAAC

402/550

FIGURE 402

MACRCLSFLLMGTFLSVSQTVLAQLDALLVFPGQVAQLSCTLSPQHVTIRDYGVSWYQQRAGS
APRYLLYYRSEEDHHRPADIPDRFSAAKDEAHNACVLTISPVPEDDADYYCSVGYGFSF

403/550

FIGURE 403

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGAAAGCAGCGAGTTG
GCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGTGGCTCACAACAAGATGCTCAAGGTGTCAGC
CGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCAGTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGG
GCGGTGCGACGGCGGTAATTTTCTGGATGATAAACAAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGG
ACAGTGGAACAAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCTTCGATCA
GGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTATGCATTGCTCAAGATTC
TCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATGAAAGAAGCAGGAGTAGACCATAGGCAGTG
GAGGGTCCCATATTATCCACCTGCAAGCAGTGCCCACTGGTCTATCCCAGCCCTGTTTGTGGTTTCAGATGGTCA
TACCTACTCTTTTCAGTGCAAACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGG
ACATTGCCCATGTCTTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCTGGAGTT
CAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAAAGTCAAAACAAGAAGACAAA
AACATTGCTGAGGCCTGAGAGAAGCAGATTGATACCAGCATCTTGCCAATTTGCAAGGACTCACTTGGCTGGAT
GTTTAACAGACTTGATACAACTATGACCTGCTATPGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAA
TGAACAGTGTACCAAGGCATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTG
CTACTGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCAAGGGGTAAA
GAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGCCAACACAATGTCATGGCAGTGT
TGGACAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTCATGGGATCCAGAATAAATGGTGTTCAGATTGTGC
TATAGATTTTGAGATCTCCGGAGATTTTGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGA
CGATATTATGAATGATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGATGA
CCATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAATCTACATTTCTAATATTTACAAAAATGATAG
CCTATTTAAATATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCACATATATTTGTATAATTATTTGAA
AAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAATPAAGAATCATTTGCTTTGAGTTTTATATTCCTTACACA
AAAAGAAAATACATATGCAGTCTAGTCAGACAAAATAAAGTTTGAAGTGCTACTATAATAAATTTTTCACGAGA
ACAACTTTGTAAATCTTCATAAGCAAAATGACAGCTAGTGTGTTGGGATCGTACATGTTAATTTTTTGAAGAT
AATTCTAAGTGAAATTTAAAATAAATAAATTTTAAATGACCTGGGTCTTAAGGATTTAGGAAAATATGCATGCT
TTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAGGATAACAGAGAGATACCACATGACTCCA
AAAAAAAAAAAAA

404/550

FIGURE 404

MLKVSAVLCVCAAAWCSQSLAAAAVAAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFRD
EVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEAGV
DHRQWRGPILSTCKQCPVVYPSPVCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCPSPDK
PTSTSRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTKTLRPERSRFDTSILPICKDS
LGWMFNRLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQRQQDP
PCQTELSNIQKRQGVKKLLGQYIPLCDEDGYKPTQCHGSVGQCWCVDRYGNEVMGSRINGVA
DCAIDFEISGDFASGDFHEWTDDEDDDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 405

[illegible]

406/550

FIGURE 406

MTPQSLQLQTTLFLLSLLFLVQGAHGRGHREDFRSCSQRNQTHRSSLHYKPTPDLRISIENSEE
ALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLLSKASSLLCFQHQE
ESLAQGPPELLATSVTSWWSPQNISLPSAASFTHSFHSPHTAAHNASVDMCELKRDLQLLSQF
LKHPQKASRRPSAAPASQQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQDLHIH
SRQEEEQSEIMEYSVLLPRTLQRTKGRSGEAEKRLLLVDFFSSQALFQDKNSSQVLGEKVLGI
VVQNTKVANLTPVVLTQHQQLQPKNVTLCVFWVEDPTLSSPGHWSSAGCETVRRETQTSCF
CNHLTYFAVLMVSSVEVDAVHKHYLSLLSYVGCVVVSALACLVTIAAYLCSRVPPLPCRRKPRDY
TIKVHMNLLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLEGYNLYRLV
VEVFGTYVPGYLLKLSAMGWGFPIFLVTLVALVDVDNYGPIILAVHRTPEGVIYPSMCWIRDS
LVSYITNLGLFSLVLEFNMAMLATMVVQILRLRPHTQKWSHVLTLGLSLVLGLPWALIFFSF
ASGTFQLVVLVLYFSIITSFQGFLIFIWYWSMRLQARGGFSPLKSNSDSARLPISSGSTSSRI

Important features:**Signal peptide:**

amino acids 1-25

Putative transmembrane domains:amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657**Microbodies C-terminal targeting signal.**

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344**G-protein coupled receptors family 2 proteins**

amino acids 475-504

407/550

FIGURE 407

TTGTGACTAAAAGCTGGCCTAGCAGGCCAGGGAGTGCAGCTGCAGGCGTGGGGGTGGCAGGAG
CCGCAGAGCCAGAGCAGACAGCCGAGAAACAGGTGGACAGTGTGAAAGAACCAGTGGTCTCGC
TCTGTTGCCCAGGCTAGAGTGTACTGGCGTGATCATAGCTCACTGCAGCCTCAGACTCCTGGA
CTTGAGAAATCCTCCTGCCTTAGCCTCCTGCATATCTGGGACTCCAGGGGTGCACTCAAGCCC
TGTTTCTTCTCCTTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTCAATCCCAAAGCTC
AGCTCTGAGCCAGAGTGGTGGTGGCTCCACCTCTGCCGCCGGCATAGAAGCCAGGAGCAGGGC
TCTCAGAAGGCCGGTGGTGGCCAGCTGGGATC**ATG**TTGTTGGCCCTGGTCTGTCTGCTCAGCTG
CCTGCTACCCCTCCAGTGAGGCCAAGCTCTACGGTCGTGTGAACTGGCCAGAGTGCTACATGA
CTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGCTTATTTAC
AAGCGGTTTCAACGCAGCTGCTTTGGACTACGAGGCTGATGGGAGCACCAACAACGGGATCTT
CCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCCGAACGTCCCCAACGTGTGCCGGAT
GTACTGCTCAGATTTGTTGAATCCTAATCTCAAGGATACCGTTATCTGTGCCATGAAGATAAC
CCAAGAGCCTCAGGGTCTGGGTACTGGGAGGCCTGGAGGCATCACTGCCAGGGAAAAGACCT
CACTGAATGGGTGGATGGCTGTGACTTCT**TAG**GATGGACGGAACCATGCACAGCAGGCTGGGAA
ATGTGGTTTGGTTCCTGACCTAGGCTTGGGAAGACAAGCCAGCGAATAAAGGATGGTTGAACG
TGAAA

408/550

FIGURE 408

MLLALVCLLSCLLPSSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALDY
EADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYSDDLNPNLKDTVICAMKITQEPQGLGYWE
AWRHHCQGKDLTEWVDGCDF

Important features:**Signal peptide:**

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

409/550

FIGURE 409

CAGACTCCAGATTTCCCTGTCAACCACGAGGAGTCCAGAGAGGAAACGCGGAGCGGAGACAACAGTACCTGACGC
 CTCTTTTCAGCCCCGGGATCGCCCCAGCAGGGATGGGCGACAAGATCTGGCTGCCCTTCCCCGTGCTCCTTCTGGCC
 GCTCTGCCTCCGGTGCTGCTGCCTGGGGCGGGCGGCTTACACCTTCCCTCGATAGCGACTTCACCTTTACCCCTT
 CCGCGCGGCCAGAAGGAGTGCTTCTACCAGCCCATGCCCTGAAGGCCTCGCTGGAGATCGAGTACCAAGTTTGA
 GATGGAGCAGGATTAGATATTGATTCCATCTTGCCTCTCCAGAAGGCCAAACCTTAGTTTTTGAACAAAGAAAA
 TCAGATGGAGTTCACACTGTAGAGACTGAAGTTGGTGATTACATGTTCTGCTTTGACAATACATTACAGCACCATT
 TCTGAGAAGGTGATTTTCTTTGAATTAATCCTGGATAATATGGGAGAACAGGCACAAGAACAAGAAGATTGGAAG
 AAATATATTACTGGCACAGATATATTGGATATGAAACTGGAAGACATCCTGGAATCCATCAACAGCATCAAGTCC
 AGACTAAGCAAAAGTGGGCACATACAAATTTCTGCTTAGAGCATTTGAAGCTCGTGATCGAAACATACAAGAAAGC
 AACTTTGATAGAGTCAATTTCTGGTCTATGGTTAATTTAGTGGTCATGGTGGTGGTGTGAGCCATTCAAGTTTAT
 ATGCTGAAGAGTCTGTTTGAAGATAAGAGGAAAAGTAGAAGTTAAACTCCAACTAGAGTACGTAACATTGAAA
 AATGAGGCATAAAAAATGCAATAAACTGTTACAGTCAAGACCATTAAATGGTCTTCTCCAAAATATTTTTGAGATATA
 AAAGTAGGAAACAGGTATAATTTTAATGTGAAAATTAAGTCTTCACTTTCTGTGCAAGTAATCCTGCTGATCCAG
 TTGTACTTAAGTGTGTAACAGGAATATTTTGCAGAAATATAGGTTTAACTGAATGAAGCCATATTAATAACTGCAT
 TTTCTTAAGTTTGAAGAAATTTTGCAGAAATGCTTAGGTGATTTAAATAAATGAGTATTGGGCCTAATTGCAACACC
 AGTCTGTTTTTAACAGGTCTTATTACCCAGAACTTTTGTGTAATGCGGCAGTTACAAATTAAGTGTGGAAGTTT
 TCAGTTTTAAGTTATAAATCACCTGAGAAATTAACCTAATGATGGATTGAATAAATCTTTAGACTACAAAAGCCCCAA
 CTTTTCTCTATTTACATATGCATCTCTCTATAATGTAAATAGAATAATAGCTTTGAAATACAATTAGGTTTTTG
 AGATTTTTATAACCAAATACATTTCACTGTAAATATAGCAGAAAGCATTAGTCTTTGTACTTTGCTTACATTC
 CCAAAAGCTGACATTTTCACGATTCTTAAAAACACAAAGTTACACTTACTAAAATTAGGACATGTTTTCTCTTTG
 AAATGAAGAATATAGTTTAAAGCTTCTCTCTCCATAGGGACACATTTTCTCTAACCCCTTAAGTAAAGTGTAGGA
 TTTTAAATTAATGTGAGGTAAATAAGTTTATTTTTAATAGTATCTGTCAAGTTAATATCTGTCAACAGTTAA
 TAATCATGTTATGTTAATTTTAAATGATTTGCTGACTTGGATAATTCATTATTACCAGCAGTTATGAAGGAAATA
 TTGCTAAATGATCTGGGCCTACCATAAATAAATATCTCTCTGAGCTCTAAGAATTATCAGAAAACAGGAA
 AGAATTTAGAAAACTTGAGAAAACCTAATCCAAATAAATTCACCTAAGTAGAAGTATAAATAAATATCTAGA
 ATCTGACTGGCTCATCATGACATCCTACTCATAACATAAATCAAAGGAGATGATTAATTTCCAGTTAGCTGGAAG
 AAATTTGGCTGTAGGTTTTATTTTCTACAAGAAATTCCTGGTTTGAATTATTTTGTAAAGCAGGTACATTTTATA
 AAATGTAAGCCCTACTGTAAGGTTTAGCACTGGGTGTACATATTTATTAATAAATTTTATTAACAACCTTTTAT
 TAAATGGCCTTTCTGAACACTTTATTTATTTGATGTTGAAGTAAGGATTAGAAACATAGACTCCCAAGTTTTAAA
 CACCTAAATGTGAATAACCCATATATACAACAAAGTTTCTGCCATCTAGCTTTTTGAAGTCTATGGGGGTCTTAC
 TCAAGTACTAGTAATTTAACTTCATCATGAATGAAGTATAATTTTTAAGTTATGCCATTTTATAACGTTGTTTAT
 GACTACATTTGTGAGTTAGAAACAACTTAAATTTGGGGTATAGAACCCCTCAACAGGTTAGTAATGCTGGAATT
 CTTGATGAGCAATAATGATAACCAGAGAGTGATTTCATTTACACTCATAGTAGTATAAAAAGAGATACATTTCCC
 TCTTAGGCCCTTGGGAGAAGAGCAGCTTAGATTTCCCTACTGGCAAGGTTTTTAAATGAGGTAAATGCCGTAT
 ATGATCAATTACCTTAATTTGGCCAAGAAATGCTTCAGGTGTCTAGGGGTATCCTCTGCAACACTTGACAGAACAA
 AGGTCAATAAGATCCTTGCCATGAATACCCCTCCCTTTTGGCGCTGTTAAATTTGCAATGAGAAGCAAAATTACA
 GTACCATAACTAATAAAGCAGGGTACAGATATAAACTACTGCATCTTTCTATAAACTGTGATTAGAATTTCTA
 CCTCTCCTGTATGGCTGTTACTGTACTGTACTCTCTGACTCCTTACCTAACAAATGAATTTGTTACATAATCTTCT
 ACATGTATGATTTGTGCCACTGATCTTAAACCTATGATTTCAGTAAGTTCTTACCATATAAAAACGATAATTGCTT
 TATTTGGAAAAGAATTTAGGAATACTAAGGACAATTTATTTTTATAGACAAAGTAAAGACAGATATTTAAGAGG
 CATAACCAAAAAGCAAACTTGTAACAGAGTAAATCTTTAATATTTCTAAAGACATACTGTTTATCTGCTT
 CATATGCTTTTTTAATTTCACTATTCCATTTCTAAATTAAGTTATGCTAAATTTGAGTAAGCTGTTTATCACTT
 AACAGCTCATTTTGTCTTTTCAATATACAAATTTTAAATACTACAAATTTAACTAAGGCCCAACCGATTTC
 CATAATGTAGCAGTTACCGTGTTCACTCACACTAAGGCCTAGAGTTTGCTCTGATATGCATTTGGATGATTAAT
 GTTATGCTGTTCTTTCACTGTGAATGTCAAGACATGGAGGGTGTGTTGTAATTTTATGGTAAATTAATCCTTCTTA
 CACATAATGGTGTCTTAAATTTGACAAAAAATGAGCACTTACAATTTGATGTCTCCTCAATGAAGATTCTTTAT
 GTGAATTTTAAAGACATTGATTTCCGCATGTAAGGATTTTTTCTCTGAAGTACAATAATGCACAATCAGTGTTG
 CTCAAACTGCTTTATACTTATAAACAGCCATCTTAAATAAGCAACGTATTGTGAGTACTGATATGTATATAATAA
 AAATTATCAAAGGAAAA

410/550

FIGURE 410

MGDKIWLPPFVLLLAALPPVLLPGAAGFTPSLDSDFTFLLPAGQKECFYQPMPLKASLEIEYQ
VLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFESTISEKVIFFELIL
DNMGEQAQEQEDWKYITGTDILDMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRNIQE
SNFDRVNFWSMVNLVVMVVVSAIQVYMLKSLFEDKRKSRT

Important features:**Signal peptide:**

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

411/550

FIGURE 411

CCCAGCTGAGGAGCCCTGCTCAAGACACGGTCACTGGATCTGAGAACTTCCCAGGGGACCGCATTCCAGAGTCA
GTGACTCTGTGAAGCACCCACATCTACCTCTTGGCCAGGTTCCCACGGGCTTGGGGGAAAGATGGTGGGGACCAAG
GCCTGGGTGTTCTCTCTCTGCTGCTGGAAGTCAACATCTGTGTTGGGGAGACAGACGATGCTCACCAGTCAGTA
AGAAGAGTCCAGCCTGGGAAGAAGAACCCAGCATCTTTGCCAAGCCTGCCGACACCCTGGAGAGCCCTGGTGAG
TGGACAACATGGTTCAACATCGACTACCCAGGCGGGAAGGGCGACTATGAGCGGCTGGACGCCATTTCGCTTCTAC
TATGGGGACCGTGTATGTGCCCGTCCCCCTGCGGGCTAGAGGCTCGGACCAGTGAAGTGGACACCTGCGGGCAGCACT
GGCCAGGTGGTCCATGGTAGTCCCCGTGAGGGTTCCTGGTGCCTCAACAGGGAGCAGCGGCCTGGCCAGAACTGC
TCTAATTACACCGTACGCTTCTCTGCCCCACCGATCCCTGCGCCGAGACACAGAGCGCATCTGGAGCCCATGG
TCTCCCTGGAGCAAGTGCTCAGCTGCCTGTGGTCAGACTGGGGTCCAGACTCGCACACGCATTTGCTTGGCAGAG
ATGGTGTGCTGTGCACTGAGGCCAGCGAAGAGGGTTCAGCACTGCATGGGCCAGGACTGTACAGCCTGTGACCTG
ACCTGCCCAATGGGCCAGGTGAATGCTGACTGTGATGCTGCCTGCATGTGCCAGGACTTCATGCTTCATGGGGCTGTC
TCCCTTCCCAGAGGTGCCCCAGCCTCAGGGGCTGCTATCTACCTCCTGACCAAGACGCCGAAGCTGCTGACCCAG
ACAGACAGTGAATGGGAGATTCCGAATCCCTGGCTTGTGCCCTGATGGCAAAGCATCCTGAAGATCACAAAGGTC
AAGTTTGGCCCCATTGTACTCACAATGCCCAAGACTAGCCTGAAGGCAGCCACCATCAAGGCAGAGTTTGTGAGG
GCAGAGACTCCATACATGGTGATGAACCCCTGAGACAAAAGCACGGAGAGCTGGGCAGAGCCTGTCTCTGTGCTGT
AAGGCCACAGGGAAGCCCAGGCCAGACAAGTATTTTGGTATCATAATGACACATTGCTGGATCCTTCCCTCTAC
AAGCATGAGAGCAAGCTGGTGTGCTGAGGAACTGCAGCAGCACCAGGCTGGGGAGTACTTTTGAAGGCCAGAGT
GATGCTGGGGCTGTGAAGTCCAAGGTTGCCAGCTGATTTGTTCACAGCATCTGATGAGACTCCTTGCAACCCAGTT
CCTGAGAGCTATCTTATCCGGCTGCCCATGATTGCTTTCAGAATGCCACCAACTCCTTCTACTATGACGTGGGA
CGCTGCCCTGTTAAGACTTGTGCAAGGCGAGCAGGATAATGGGATCAGGTGCCGTGATGCTGTGCAAGAACTGCTGT
GGCATCTCAAGACAGAGGAAGGGAGATCCAGTGCACTGGCTACACGCTACCCACCAAGGTGGCCAGGAGGATGC
AGCTGCCAGCGGTGTACGGAACTCGGAGCACTCGTGGGGGCGCTGTGCTGCTGCTGCTGCAATGGGGAGCCCATG
CGCTTTGGCCATGTGTACATGGGGAAACAGCCGTGTAAGCATGACTGGCTACAAGGGCACTTTACCCCTCCATGTC
CCCCAGGACACTGAGAGGCTGGTGTCACTTTGTGGACAGGCTGCAGAASTTTGTCAACACCACCAAGTGGTA
CCTTTCAACAAGAAGGGGAGTGCCGTGTTCCATGAAATCAAGATGCTTCGTGCGAAAGAGCCCATCACTTTGGAA
GCCATGGAGACCAACATCATCCCCCTGGGGGAAGTGGTTGGTGAAGACCCCATGGCTGAAGTGGAGATTCCATCC
AGGAGTTTCTACAGGCAGAAATGGGGAGCCCTACATAGGAAAGTGAAGGCCAGTGTGACCTTCTGGATCCCCGG
AATATTTCCACAGCCACAGCTGCCAGACTGACCTGAACTTCATCAATGACGAAGGAGACACTTTCCCCCTTCGG
ACGTATGGCATGTTCTCTGTGGACTTCAGAGATGAGGTCACCTCAGAGCCACTTAATGCTGGCAAAGTGAAGGTC
CACCTTGACTCGAACCAGGTCAAGATGCCAGAGCACATATCCACAGTGAAACTCTGGTCACTCAATCCAGACACA
GGGCTGTGGGAGGAGGAAGGTGATTTCAAAATTTGAAATCAAAGGAGGAACAAAAGAGAAGACAGAACTTCTCTG
GTGGGCAACCTGGAGATTCTGTGAGAGGAGGCTCTTAACCTGGATGTTCTCTGAAAGCAGGCGGGTCTTTGTTAAG
GTGAGGGCCTACCGGAGTGAGAGGTTCTTGCTAGTGAGCAGATCCAGGGGGTTGTGATCTCCGTGATTAACCTG
GAGCCTAGAAGTGGCTTCTGTCCAACCCCTAGGGCCTGGGGCCGCTTTGACAGTGTATCACAGGCCCAACGGG
GCCTGTGTGCTGCTTCTGTGATGACCACTCCCTGATGCTTACTCTGCCTATGCTCTTGGCAAGCCTGGCTGGG
GAGGAAGTGAAGCAGTGGAGTCTTCTCTAAATTCACCCAAATGCAATTTGGCGTCCCTCAGCCCTATCTCAAC
AAGCTCAACTACCGTCGGACGGACCATGAGGATCCACGGGTTAAGAAAGACAGCTTTCCAGATTAGCATGGCCAAG
CCAAGGCCCAACTCAGCTGAGGAGAGCAATGGGCGCTATGCTGCTTTGAGAACCTCCGGGATGTGAAGAGGCA
CCACCCAGTGACGCCCACTTCCGGTCTTACAGATGAGGGGGATCGATATGACTACAACACAGTCCCTTCAAC
GAAGATGACCCATGAGCTGGACTGAAGACTATCTGGCATGGTGGCCAAAGCCGATGGAATTCAGGGCCTGCTAT
ATCAAGGTGAAGATTGTGGGGCCACTGGAAGTGAATGTGCGATCCCGCAACATGGGGGGCACTCATCGGCGGACA
GTGGGGAAGCTGTATGGAATCCGAGATGTGAGGAGCACTCGGGACAGGGACCAGCCCAATGTCTCAGCTGCCTGT
CTGGAGTTCAAGTGCAGTGGGATGCTCTATGATCAGGACCGTGTGGACCGCACCCCTGGTGAAGGTATCCCCAG
GGCAGCTGCCGTGAGCCAGTGTGAACCCCATGCTGCATGAGTACCTGGTCAACCACTTGGCACTTGCAGTCAAC
AACGACACCAAGTGAGTACACCATGCTGGCACCCCTTGGACCCCACTGGGGCCACAACATGCGCATCTACACTGTCACT
GACCAGGACCCCTCGCACGGCCAAGGAGATCGCGCTCGGGCGGTGCCTTGTATGGCACATCCGATGGCTCCTCCAGA
ATCATGAAGAGCAATGTGGGAGTAGCCCTCACCTTCAACTGTGTAGAGAGGCAAGTAGGCCGCCAGAGTGCCTTC
CAGTACCTCCAAAGCACCCAGCCAGTCCCTGCTGCAGGCACTGTCCAAGGAAGAGTGGCCTCGAGGAGGCAG
CAGCGAGCGAGCAGGGGTGGCCAGCGCCAGGGTGGAGTGGTGGCTCTCTGAGATTTCTAGAGTTGCTCAACAG
CCCCTGATCAACTAAGTTTTGTGGTACTTCACCTCTTCTGCCCTCATTTTCATGTGACAGCCATTGTGAGACTGA
TGCACAACTGTCACTTGGTTAATTTAAGCACTTCTGTTTTCTGTAATTTGCTTGTGTTTCTTTCATGCCTTTA
CTTACTTTGTCCCATGCTACTGATTGGCACGTGGGCCCCACAATGGCACAATAAAGCCCTTTGTGAACTGTTT
TTTAAATGAAACACAAGAAATTTGGCCACTGGTAAACTCTGCAGCTTCAACTGTACTTCATTTAATGCCATTAAT
GCAATATACTTCTCTTTTGTGATGGTPTTGGCCACCTCTGCAATAGTGATAATCTGATGCTGAAGATCAA
ATAACCAATATAAGCATATTTCTTGGCCTTGTCCACAGGCAATAGGCAAGCCTTGATCATAGTTTCATACATAT
AATGGTGGTGAATAAAGAAATAAACAACAATACTTTTACTTGAATGTAAATACTTATTTATTTCTTTGCTA
AATTTGGAATTTAGTGACATTTCAAAGTTAAGCTATTAATATAGGGTGATCATAGTTCTCTACCAAGTCTGG
AAAGAACATCTCTGGTATCCACAATTACACCAGGTTGCTAACTGTATTTGTACATTTCCCTTTGCATTGCTTT
TGTTCTTGCTAGAAACCAAGTGTAGCCAGGGCAGATGTCAATAAATGCATFACTCTGTATTTGCAAAAAA

412/550

FIGURE 412

MVGTKAWVFSFLVLEVTSVLGRQTMLTQSVRRVQPGKKNPSIFAKPADTLES PGWTTWFNID
YPGGKGDYERLDAIRFYYGDRVCARPLRLEARTTDWTPAGSTGQVVHGSPREGFWCLNREQRP
GQNC SNYTVRFLCPPGSLRRDTERIWSFWSKCSAACGQTGVQTRTRICLAEMVSLCSEAS
EEGQHCMGQDCTACDLTCFPMGQVNADCDACMCQDFMLHGAVSLPGGAPASGAAYLLTKTPKL
LTQTDS DGRFRIPGLCPDGKSILKITKVKEAPIVLTMPKTS LKAATIKAEFVRAETPYMVMNP
ETKARRAGQSVSLCCKATGKPRPDKYFWYHNDTLLDPSLYKHESKLVLRKLQHQAGEYFCKA
QSDAGAVKSKVAQLIVTASDETFCNFPVPE SYLIRLPHDCFQNATNSFYDVGRCVPKTCAGQQ
DNGIRCRDAVQNCCGISKTEEREIQCSGYTLPTKVAKESCQRCTETRSIVRGRVSAADNGEP
MRFGHVYMGNSRVSM TGYKGTFTLHVPQDTERLVLT FVDRLQKEVNTTKVLPENKKGSAVFHE
IKMLRRKEPITLEAMETNIIPLGEVVGEDEMAELEIPSR SFYRQNGEPYIGKVKASVTFLDPR
NISTATAAQTDLNF INDEGDTFPLRTYGMFSVDFRDEVTSEPLNAGKVKVHLDSTQVKMPEHI
STVKLWSLNPDTGLWEEEGDFKFENQRRNKREDRTFLVGNLEIRERRLENLDVPESRRCFVKV
RAYRSE RFLPSEQIQGVVISVINLEPRTGFLSNPRAWGRFDSVITGPNGACVP AFCDQSPDA
YSAYVLASLAGEELQAVESSPKFNPNAIGVPQPYLNKLN YRRTDHEDPRVKKTAFQISM AKPR
PNSAEESNGPIYAFENLRACEEAPPSAAHFRFYQIEGDRYDYNTVPFNEDDPM SWTEDYLAWW
PKPMEFRACYIKVKIVGPLEVNVR SRNMGGTHRRTVGKLYGIRDVRSTRDRDQPNVSAACLEF
KCSGMLYDQDRVDRTL VKVIPQGSCRRASVNPMLHEYLVNHLPLAVNNDTSEY TMLAPLDPLG
HNYGIYTVTDQDPRTAKEIALGRCFDGTSDGSSRIMKSNVGVALTFNCVERQVGRQSAFQYLO
STPAQSPAAGTVQGRVPSRRQQRASRGGQRQGGVVASLRFFRVAQQPLIN

413/550

FIGURE 413

GCCACGTTGTCTTCTTTCCTTCACCACCACCCAGGAGCTCAGAGATCTAAGCTGCTTTCCATC
TTTTCTCCCAGCCCCAGGACACTGACTCTGTACAGGAATGGGGCCGTCCTCTTGCCCTCCTTCTC
ATCCTAATCCCCCTTCTCCAGCTGATCAACCCGGGGAGTACTCAGTGTTCCCTTAGACTCCGTT
ATGGATAAGAAGATCAAGGATGTTCTCAACAGTCTAGAGTACAGTCCCTCTCCTATAAGCAAG
AAGCTCTCGTGTGCTAGTGTCAAAAGCCAAGGCAGACCGTCCTCCTGCCCTGCTGGGATGGCT
GTCAGTGGCTGTGCTTGTGGCTATGGCTGTGGTTCGTGGGATGTTTCAGCTGGAAACCACCTGC
CACTGCCAGTGCAGTGTGGTGGACTGGACCACTGCCCCGCTGCTGCCACCTGACCTTGACAGGGA
GGAGGCTGAGAACTCAGTTTTGTGACCATGACAGTAATGAAACCAGGGTCCCAACCAAGAAAT
CTAACTCAAACGTCCCACTTCATTTGTTCCATTCCTGATTCTTGGGTAATAAAGACAAACTTT
GTACCTCAAAAAAAAAAAAAAAAAAAAAA

414/550

FIGURE 414

MGPSSCLLLILIPLLQLINPGSTQCSLDSVMDKKIKDVLNSLEYSPSPISKKLSCASVKS
QGRPSSCPAGMAVTGCACGYGCGSWDVQLETTCHCQCSVVDWTTARCCHLT

415/550

FIGURE 415

CAGAAGAGGGGGCTAGCTAGCTGTCTCTGCGGACCAGGGAGACCCCCGCGCCCCCCCCGGTGTG
AGGCGGCCTCACAGGGCCGGGTGGGCTGGCGAGCCGACGCGGCGGGCGGAGGAGGCTGTGAGGA
GTGTGTGGAACAGGACCCGGGACAGAGGAACCATGGCTCCGCAGAACCTGAGCACCTTTTGCC
TGTGCTGCTATACCTCATCGGGGCGGTGATTGCCGGACGAGATTTCTATAAGATCTTGGGGG
TGCCTCGAAGTGCCTCTATAAAGGATATTAAAAAGGCCTATAGGAAACTAGCCCTGCAGCTTC
ATCCCGACCGGAACCCTGATGATCCACAAGCCCAGGAGAAATTCCAGGATCTGGGTGCTGCTT
ATGAGGTTCTGTCAGATAGTGAGAAACGGAAACAGTACGATACTTATGGTGAAGAAGGATTAA
AAGATGGTCATCAGAGCTCCCATGGAGACATTTTTTCACACTTCTTTGGGGATTTTGGTTTCA
TGTTTGGAGGAACCCCTCGTCAGCAAGACAGAAATATTCOAAGAGGAAGTGATATTATTGTAG
ATCTAGAAGTCACTTTGGAAGAAGTATATGCAGGAAATTTTGTGGAAGTAGTTAGAAACAAAC
CTGTGGCAAGGCAGGCTCCTGGCAAACGGAAGTGCAATTGTGCGCAAGAGATGCGGACCACCC
AGCTGGGGCCCTGGGCGCTTCCAAATGACCCAGGAGGTGGTCTGCGACGAATGCCCTAATGTCA
AACTAGTGAATGAAGAACGAACGCTGGAAGTAGAAATAGAGCCTGGGGTGAGAGACGGCATGG
AGTACCCCTTTATTGGAGAAGGTGAGCCTCACGTGGATGGGGAGCCTGGAGATTTACGGTTCC
GAATCAAAGTTGTCAAGCACCCAATATTTGAAAGGAGAGGAGATGATTTGTACACAAATGTGA
CAATCTCATTAGTTGAGTCACTGGTTGGCTTTGAGATGGATATTACTCACTTGGATGGTCACA
AGGTACATATTTCCCGGGATAAGATCACCAGGCCAGGAGCGAAGCTATGGAAGAAAGGGGAAG
GGCTCCCCAACTTTGACAACAACAATATCAAGGGCTCTTTGATAATCACTTTTGATGTGGATT
TTCCAAAAGAACAGTTAACAGAGGAAGCGAGAGAAGGTATCAAACAGCTACTGAAACAAGGGT
CAGTGCAGAAGGTATACAATGGACTGCAAGGATATTGAGAGTGAATAAAATTGGACTTTGTTT
AAAATAAGTGAATAAGCGATATTTATTATCTGCAAGGTTTTTTTGTGTGTGTTTTTGTTTT
TTTTCAATATGCAAGTTAGGCTTAATTTTTTTATCTAATGATCATCATGAAATGAATAAGAGG
GCTTAAGAATTTGTCCATTTGCATTCGGAAAAGAATGACCAGCAAAAGGTTTACTAATACCTC
TCCCTTTGGGGATTTAATGTCTGGTGCTGCCGCCTGAGTTTCAAGAATTAAAGCTGCAAGAGG
ACTCCAGGAGCAAAAGAAACACAATATAGAGGGTTGGAGTTGTTAGCAATTTCAATCAAATG
CCAACTGGAGAAGTCTGTTTTTAAATACATTTTGTTGTTATTTT

416/550

FIGURE 416

MAPQNLSTFCLLLLLYLIGAVIAGRDFYKILGVPRSASIKDIKKAYRKALQLHPDRNPDDPQAQEKFDLGAAYE
VLSDSEKRKQYDTYGEEGLKDGHQSSHGDI FSHFFGDFGFMFGGTPRQQDRNIPRGSDIIVDLEVTLEEVYAGNF
VEVVRNKPVARQAPGKRKCNCRQEMRTTQLGPGRFQMTQEVVCDECPNVKLVNEERTLEVEIEPGVRDGMETPFI
GEGEPHVDGEPDGLRFRIKVVVKHPIFERRGDDLYTNVTISLVESLVGFEMDITHLDGHKVVHISRDKITRPGAKLW
KKGEGLPNFDNNNIKGLIITFDVDFPKEQLTEEAREGIKQLLKQGSVQKVYNGLQGY

Important features:**Signal peptide:**

amino acids 1-22

Cell attachment sequence.

amino acids 254-257

Nt-dnaJ domain signature.

amino acids 67-87

Homologous region to Nt-dnaJ domain proteins.

amino acids 26-58

N-glycosylation site.

amino acids 5-9, 261-265

Tyrosine kinase phosphorylation site.

amino acids 253-260

N-myristoylation site.

amino acids 18-24, 31-37, 93-99, 215-221

Amidation site.

amino acids 164-168

417/550

FIGURE 417

CGGCGGCGGCTGCGGGCGCGAGGTGAGGGGCGCGAGGTGAGGGGCGCGAGGTTCCCAGCAGGA
TGCCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCCGGAGAGGGCCCAGCCCGCCCGGGGCAG
GATGACCAAGGCCCGGCTGTTCCGGCTGTGGCTGGTGCTGGGGTCGGTGTTCATGATCCTGCT
GATCATCGTGTACTGGGACAGCGCAGGCGCCGCGCACTTCTACTTGACACAGTCCTTCTCTAG
GCCGCACACGGGGCCGCGCTGCCACGCCCCGGGCCGACAGGGACAGGGAGCTCACGGCCGA
CTCCGATGTCGACGAGTTTCTGGACAAGTTTCTCAGTGCTGGCGTGAAGCAGAGCGACCTTCC
CAGAAAGGAGACGGAGCAGCCGCCTGCGCCGGGGAGCATGGAGGAGAGCGTGAGAGGCTACGA
CTGGTCCCCGCGCGACGCCCCGGCGCAGCCCAGACCAGGGCCGGCAGCAGGCGGAGCGGAGGAG
CGTGCTGCGGGGCTTCTGCGCCAACTCCAGCCTGGCCTTCCCCACCAAGGAGCGCGCATTCGA
CGACATCCCCAACTCGGAGCTGAGCCACCTGATCGTGGACGACCGGCACGGGGCCATCTACTG
CTACGTGCCCAAGGTGGCCTGCACCAACTGGAAGCGCGTGATGATCGTGCTGAGCGGAAGCCT
GCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGCGAGCACGTGCACAACGC
CAGCGCGCACCTGACCTTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCTCCCGCCACCTCAT
GAAGGTCAAGCTCAAGAAGTACACCAAGTTCCTCTTCGTGCGCGACCCCTTCGTGCGCCTGAT
CTCCGCCTTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCGCAAGTTGCGCGTGCC
CATGCTGCGGCTGTACGCCAACCACACCAGCCTGCCCGCCTCGGCGCGCGAGGCCCTTCCGCGC
TGGCCTCAAGGTGTCCTTCGCCAACTTCATCCAGTACCTGCTGGACCCGCACACGGAGAAGCT
GGCGCCCTTCAACGAGCACTGGCGGCAGGTGTACCGCCTCTGCCACCCGTGCCAGATCGACTA
CGACTTCGTGGGGAAGCTGGAGACTCTGGACGAGGACGCCGCGCAGCTGCTGCAGCTACTCCA
GGTGGACCGGCAGCTCCGCTTCCCCCGAGCTACCGGAACAGGACCGCCAGCAGCTGGGAGGA
GGACTGGTTCGCCAAGATCCCCTGGCCTGGAGGCAGCAGCTGTATAAACTCTACGAGGCCGA
CTTTGTTCTCTTCGGCTACCCCAAGCCCGAAAACCTCCTCCGAGACTTGAAAGCTTTCGCGTTG
CTTTTTCTCGCGTGCTGGAACCTGACGCACGCGCACTCCAGTTTTTTTTATGACCTACGATTT
TGCAATCTGGGCTTCTTGTTCACTCCACTGCCTCTATCCATTGAGTACTGTATCGATATTGTT
TTTAAAGATTAAATATATTTTCAGGTATTTAATACGA